Smoothing of probabilities of death for older people in life expectancy tables

Gustaf Strandell and Tomas Johansson[1]

[1]Statistics Sweden, Örebro, Sweden

Email: gustaf.strandell@scb.se

Email: tomas.johansson@scb.se

Abstract: This paper focuses on the smoothing procedure used at Statistics Sweden for handling of probabilities of death for persons in the highest ages, where the population is small and mortality is high. The paper also demonstrates how the smoothing affects the estimated average life expectancies at the national and regional levels.

Keywords: Life table, Probabilities of death, Smoothing, Life expectancy

1 Introduction

Life expectancy tables in Sweden are based on data on population size and deaths during either a one-year or a five-year period. The tables are produced annually for the whole country, for its 21 counties and 290 municipalities (the table based on one year of data is only produced for the whole country). In addition to the remaining average life expectancies, the publication of the life expectancy table includes sex-specific probabilities of death for all ages.

In 2012, Statistics Sweden conducted a review of the calculations for the periodic life expectancy tables. The project had two main objectives, of which the second is addressed in this report:

1. A quality assured production system, and
2. Review of the handling of probabilities of death for very old persons at the national and regional levels.

Usable estimates of probabilities of death q_x can be difficult to obtain for the highest ages in a population. Since there might be few survivors and hence few deaths in higher ages the observed probabilities of death can often be seen to fluctuate significantly between ages or between time periods.

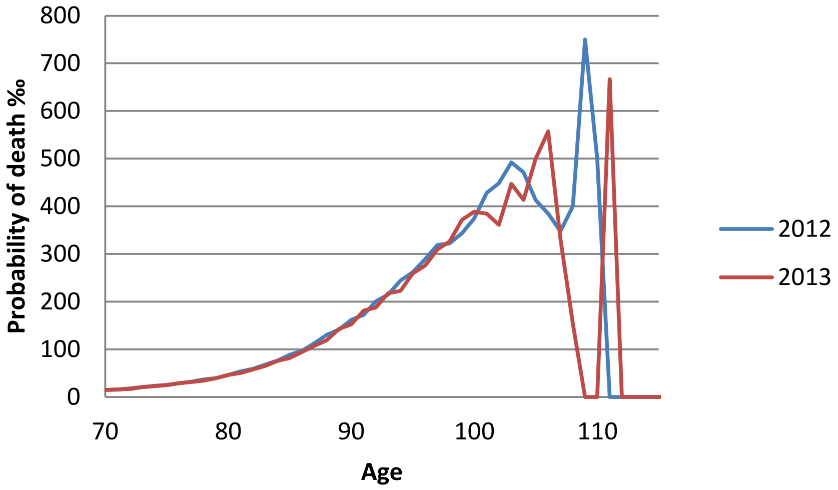

Fig 1. Non smoothed probabilities of death (‰) for the population of Sweden 2012 and 2013

To address this problem Statistics Sweden applies a smoothing procedure to the probabilities of death for the highest ages and over the years various methods to do this has been tried. Up to 1986 a method due to Wittstein was used, which overestimated the probabilities of death for the oldest persons. Therefore in 1987, Sten Martinelle at Statistics Sweden constructed a new smoothing procedure which was used successfully for many years. However, this method, where by Martinelles modeled probabilities of death replaced the observed probabilities for ages 90 and above, relied heavily on mortality data for Sweden and some other countries up to 1987 and did not fully take into account the observed probabilities to be smoothed. Therefore, over time, the smoothing process deteriorated, resulting in a systematic underestimation of probabilities of death and an overestimation of the remaining life expectancy for newborns, e_0. Thus, in 2012, when the calculations for the life expectancy table were up for review it was decided that Martinells method was to be evaluated, resulting in an updated method which provides smoothed probabilities of death with a better fit to the observed data.

This report focuses on the updated method used since 2012 by Statistics Sweden. It also demonstrates how the updated smoothing affects estimated average life expectancies at the national and regional levels.

2 The model for smoothing probabilities of death

The updated method used by Statistics Sweden for smoothing of the probabilities of death in high ages is still based on the model chosen by Sten Martinell for old-age mortality, Martinelle (1987).

The basic assumption underlying the model states that age specific adult mortality for a single individual in a population, in terms of the force of mortality $\mu(x)$ where x denotes age, is well approximated by the Gompertz-Makeham formula

$$\mu(x) = A + Be^{kx} \qquad A \geq 0, B > 0, k > 0.$$

The constant B is called the 'frailty' of the individual and measures his or her inability to withstand destruction.

In the past many authors has pointed out that it is medically well established that frailty is different for different individuals in a population. It is hence customary to assume that frailty follows a probabilistic distribution over a population rather than being the same number for all individuals. It was proved by Beard (1959) that if the frailty B is gamma distributed then the mean of the force of mortality taken over the population follows the Perks formula, as suggested by Perks already in 1932, Perks (1932):

$$\mu(x)=\frac{A+Be^{kx}}{1+De^{kx}}\qquad A\geq 0, B>0, D>0, k>0$$

Martinelles main objection towards using Perks formula is that it implies that the force of mortality approaches a constant limit value B/D in high ages. He argues that there is no empirical evidence for a plateau in the mortality for centenarians. Therefore he modifies the assumption leading to Perks formula, and replaces the assumption of a gamma distributed frailty with an assumption of a frailty with a generalized gamma distribution which has been shifted to the right, thus assuming that it is impossible for the frailty to assume values very close to zero.

In the case where the constant A in the Gompertz-Makeham formula is assumed to be zero, which seems reasonable in Sweden in modern times, Martinelles model of the frangible man states the following:

If the force of mortality for individuals with frailty Z follow the Gompertz-Makeham law $\mu(x|z)=ze^{kx}$ and the frailty variable Z has a shifted gamma distribution with density

$$g(t)=\begin{cases}\dfrac{b^a(t-c)^{a-1}}{\Gamma(a)}e^{-b(t-c)} & \text{for } t>c\\ 0 & \text{for } t\leq c\end{cases}$$

($a,b,c>0$), then the mean force of mortality over the population, $\mu(x)$, is given by

$$\mu(x)=\left(c+\frac{\eta}{1+\eta\alpha^2\int_{x_0}^{x}e^{kt}dt}\right)e^{kx}$$

Where η and α are the mean and the relative standard deviation of the variable $Z-c$.

As usual the probabilities of death q_x is associated with the force of mortality by $q_x = 1 - e^{-\mu(x)}$. Substituting the expression above for the modelled mean force of mortality into this expression for q_x, we end up with the formula we use for smoothing of the probabilities of death for high ages

$$q_x = 1 - e^{-\left(c + \frac{\eta}{1 + \eta\alpha^2 \int_{x_0}^{x} e^{kt} dt}\right) e^{kx}} .$$

This far, the new updated procedure used by Statistics Sweden since 2012 follows Martinells suggestions in Martinelle (1987). The fitting of the model to mortality data is however new and improved. Instead of using a large set of historical data the model is now, whenever possible, fitted using only probabilities of death observed for the region and the time period for which the life expectancy table is to be produced. Thus the smoothed probabilities now more explicitly take into account the observed difference in mortality between regions and between time periods.

The model for the age specific probabilities of death contains 4 parameters: α, k, c and η. The modeled probabilities dependence on α and k is indicated in the diagram below where $c = 1{,}363 \cdot 10^{-7}$ and $\eta = 5{,}535 \cdot 10^{-7}$.

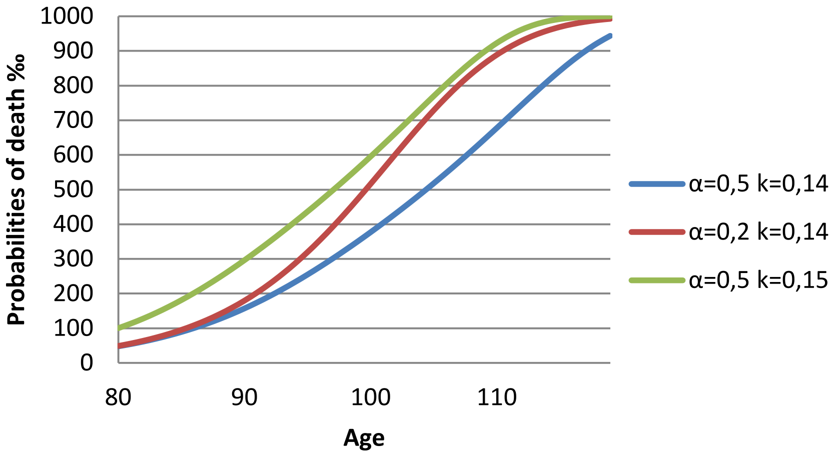

Fig2. Modeled probabilities of death

For the very high ages a decrease in α leads to a steeper increase of the probabilities of death between ages, an increase in k leads to an almost constant increase of the probabilities for a broad interval of high ages.

It is technically cumbersome to fit this model to mortality data without first fixing two of the parameters. After testing numerous different parameter settings against data we decided to fix the parameters $\alpha = 0{,}5$ and $k = 0{,}14$. Our experimentation also showed that it is not recommended to use probabilities of death for all available ages to fit the model. We got much better results when we restricted the mortality data used as input to the fitting procedure to ages in a span between m and n, to be specified below, then when we used all ages or even all higher ages. Finally we decided it preferable to use age specific weights in the fitting of the model. The reason for this is that especially for smaller populations, the probabilities of death for certain ages can be zero (if no one dies) ore one (if everybody dies) and these extreme values has a tendency to gain an unjustified influence on the modeled probabilities of death for the other ages. Therefore we decided to use the number of deaths per age as weights in the fitting of the model.

To fit the model to observed probabilities of death we use a weighted least square method: If the model $q_x = q_x(\zeta,\eta)$, where α and κ has been fixed as above, is to be fitted to the observed probabilities of death $\tilde{q}_m, \ldots, \tilde{q}_n$ over the age interval from age m to age n, we try to find ζ and η which minimizes the expression

$$\frac{1}{n-m+1}\sum_{x=m}^{n} v_x^2 \left(\tilde{q}_x - q_x(\zeta,\eta)\right)^2$$

where v_x are the nonnegative weights. To do this in practice we use the *model procedure* in SAS.

3 Results

At the national level, the model is fitted using probabilities of death observed in the age interval between $m=90$ and $n=100$ years. Since the number of individuals in the higher ages, and hence the number of deaths among them, has increased over the years the age at which the smoothing start to take place was also increased from 91 to 95 years.

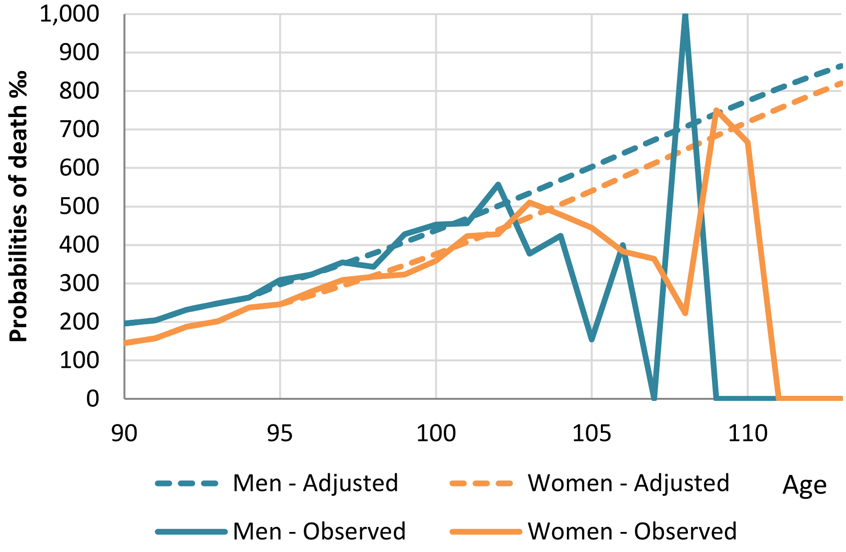

Fig 3. Observed and adjusted probabilities of death by sex and age (90 to 113) for Sweden 2012

The effect of the update of the smoothing method on the probabilities of death can be seen in the diagram below. For example the probability of death for women age 100 in 2011 was raised by over 10% with the updated method compared with the old one.

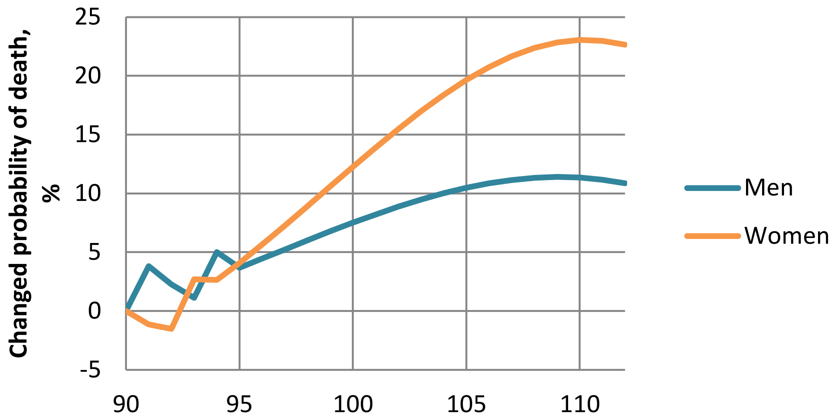

Fig 4. Difference in probability of death between new and old smoothing method 2011

Although the probabilities of death for ages 95 and over have clearly changed with the new method compared with the old one, the update of the smoothing method did not bring about particularly large changes in average life expectancies for the whole of Sweden. As suspected, the updated method gives life expectancies which are very close to those gained if the life expectancy table were to be produced without any smoothing. For the whole country the smoothing has thus more of a cosmetic effect on the mortality data (which is a good thing and can be very important for the presentation of the data).

Sex	**Age**	**No smoothing**	**Old smoothing method**	**Updated smoothing method**
Male	**e0**	79.8	79.81	79.79
	e85	5.51	5.54	5.5
Female	**e0**	83.67	83.7	83.67
	e85	6.61	6.66	6.61

Fig 5. Average life expectancies for Sweden 2011

At the county level, in order to capture regional differences in mortality among the elderly, the model is also fitted using the observed probabilities of death. However, the population in most counties is too small to appropriately apply the same method for the counties as for the national level. After extensive testing, it was decided that the fitting of the model would be based on observations in the age interval of 80 to 100. It was also decided that the smoothed probabilities of death would be used from the age of 90.

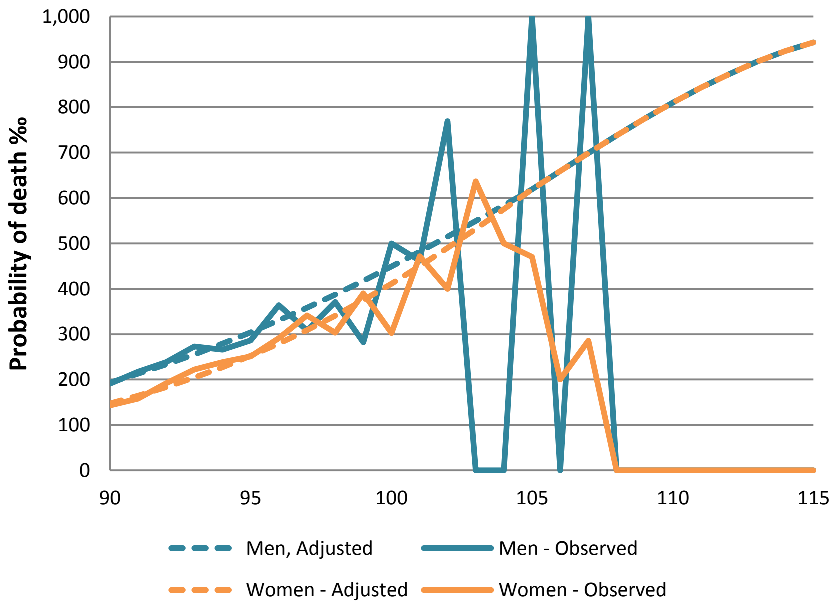

Fig 6. Observed and adjusted probabilities of death (‰) by sex and age (90 to 115) for Örebro county 2012

The update of the smoothing method has a larger impact on the life expectancies for the counties than for the whole of Sweden, but the effects are not dramatic. Among the counties we noted the greatest change for the county of Norrbotten (2007-2011) where the updated method subtracts about a month from the mean total lifespan of women. Also for the counties, the updated method (compared with the old one) gives life expectancies which are closer to those gained if the life expectancy table were to be produced without any smoothing.

In the smoothing of the probabilities of death for municipalities, the population and hence the number of deaths are in many cases too small to use as basis for the fitting of the model. There are 290 municipalities In Sweden with populations ranging from just under 2 500 inhabitants to nearly 900 000 inhabitants. One-half of the municipalities have a population of around 15 000 inhabitants or less. Therefore, the smoothed probabilities of death at the corresponding county level are used for all the constituent municipalities from the age of 90. An exception is made for the three largest municipalities: Stockholm, Gothenburg and Malmö, where the adjustment of probabilities of death is based on the individual municipalities in the same way as for the counties.

The updated method has had a greater impact at the municipal level than at the national and county levels. Based on the period 2007-2011, the average life expectancy for newborn females is affected up to a shorter life expectancy of just over three months, compared with the old smoothing procedure. The largest increase in average life expectancy for newborns in a municipality was an increase of about two months.

At the national level and for some of the counties the effect of the smoothing on the average life expectancies is marginal, compared to results gained without smoothing. However, for many municipalities the smoothing of the probabilities of death is absolutely necessary. Without it, for some municipalities, we would end up with very unstable average life expectancies. For example, for the small municipality of Jokkmokk in the north of Sweden, the smoothing subtracts 150 days from e_0 for women.

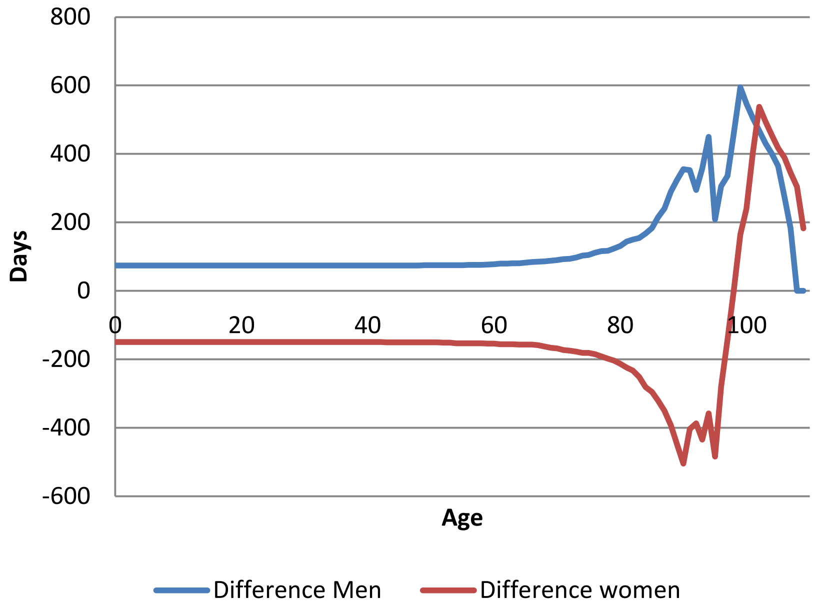

Fig 7. Difference in average life expectancy by age between smoothed and not smoothed results for Jokkmokk 2013

References

Beard, R. E. (1959): Note on Some Mathematical Mortality Models, Colloquia on Ageing, CIBA Foundation, Vol 5. Churchill, London

Martinelle, M. (1987): A Generalized Perks Formula for Old-Age Mortality, R&D Report, U/STM-39, Statistics Sweden

Perks, W. (1932): On some Experiments in the Graduation of Mortality Statistics. Journal of the Institute of Actuaries. Vol LXIII, 1932, p12-57

Statistics Sweden, Slutrapport 2012 BVBE Livslängdstabeller (2012)